

## **Pendiskretan Data Set Kasar Menggunakan Ta'akulan Boolean**

**ROKIAH @ ROZITA AHMAD, MASLINA DARUS, SITI MARIYAM SHAMSUDDIN  
& AZURALIZA ABU BAKAR**

### **ABSTRAK**

*Pendiskretan data set kasar terhadap nilai nyata atribut merupakan salah satu aspek yang paling penting di dalam konsep pelombongan data, terutamanya yang melibatkan masalah pengkelasan masalah. Hasil kajian secara empirik menunjukkan bahawa kualiti pengkelasan bergantung kepada algoritma pendiskretan yang digunakan di dalam fasa pra-pemprosesan input data. Oleh yang sedemikian, proses pendiskretan boleh dinyatakan sebagai satu operasi gelidahan partisi terhadap domain atribut ke dalam bentuk selang dan menggabungkan nilai tersebut ke atas setiap selang yang berkaitan. Proses pendiskretan melibatkan gelidahan terhadap potongan yang menentukan selang berkaitan. Kesemua nilai yang berada pada setiap selang tersebut dipetakan terhadap nilai yang sama, iaitu menukarkan nilai atribut berkaitan kepada satu nilai berangka, yang boleh dijadikan sebagai atribut simbolik. Gelidahan terhadap potongan dilaksanakan terhadap perwakilan dalaman integer bagi input sistem keputusan. Kertas kerja ini bertujuan mengupas peranan set kasar dalam pendiskretan data menggunakan ta'akulan Boolean yang melakukan proses penukaran nilai nyata simbol matematik bercetak ke dalam perwakilan yang lebih mudah untuk menghasilkan kadar pencaman yang lebih baik menggunakan rangkaian nural.*

### **ABSTRACT**

*Data discretization of rough set towards real attribute values is one of the important aspect in the data mining concepts, particularly the ones which involved classification problems. Empirical results showed that the quality of the classification depends on the discretization algorithm used in the input data pre-processing phase. In general, discretization is a process of search-ing for partition of attribute domains into intervals and unifying the values over each interval. Discretization process involves searching of cuts which determine the intervals acquired. All values which lie within each interval are mapped to the same values, in effect converting numerical attributes that can be treated as being symbolic. The search for cuts is performed on the internal integer representation of the input decision system. This paper describes the role of rough set with Boolean reasoning discretization process in converting the real values of printed mathematical symbol that leads to the better recognition rates using neural network.*